



**Barcelona  
Supercomputing  
Center**

*Centro Nacional de Supercomputación*



EXCELENCIA  
SEVERO  
OCHOA

# Adaptive Ensemble Climate Simulations

Alicia Sanchez Lorente on behalf of the Earth Sciences  
Department at BSC



- Earth activities at BSC
- Ensemble Simulation for uncertainty characterization
  - EC-Earth climate model
  - Adaptive ensemble workflow
- Discussion



# BSC Earth Department

## What

Environmental forecasting

## Why

Our strength ...  
research  
operations  
services  
more than 60 people working together

## How

Develop a capability to model air quality processes from urban to global and the impacts on weather, health and ecosystems

Implement a climate prediction system for subseasonal-to-decadal climate prediction

Develop user-oriented services that favour both technology transfer and adaptation

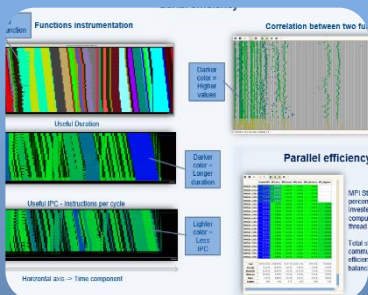
Use cutting-edge HPC and Big Data technologies for the efficiency and user-friendliness of Earth system models

Earth system  
services

Climate  
prediction

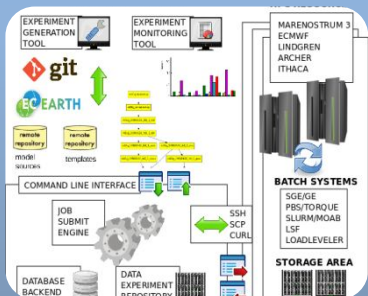
Atmospheric  
composition

Computational  
Earth sciences



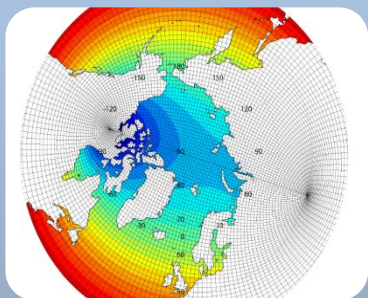
## Performance Team

- Provide HPC Services
- Apply new computational methods



## Models and Workflows Team

- Development of HPC user-friendly software framework
- Support the development of climate and atmospheric research software

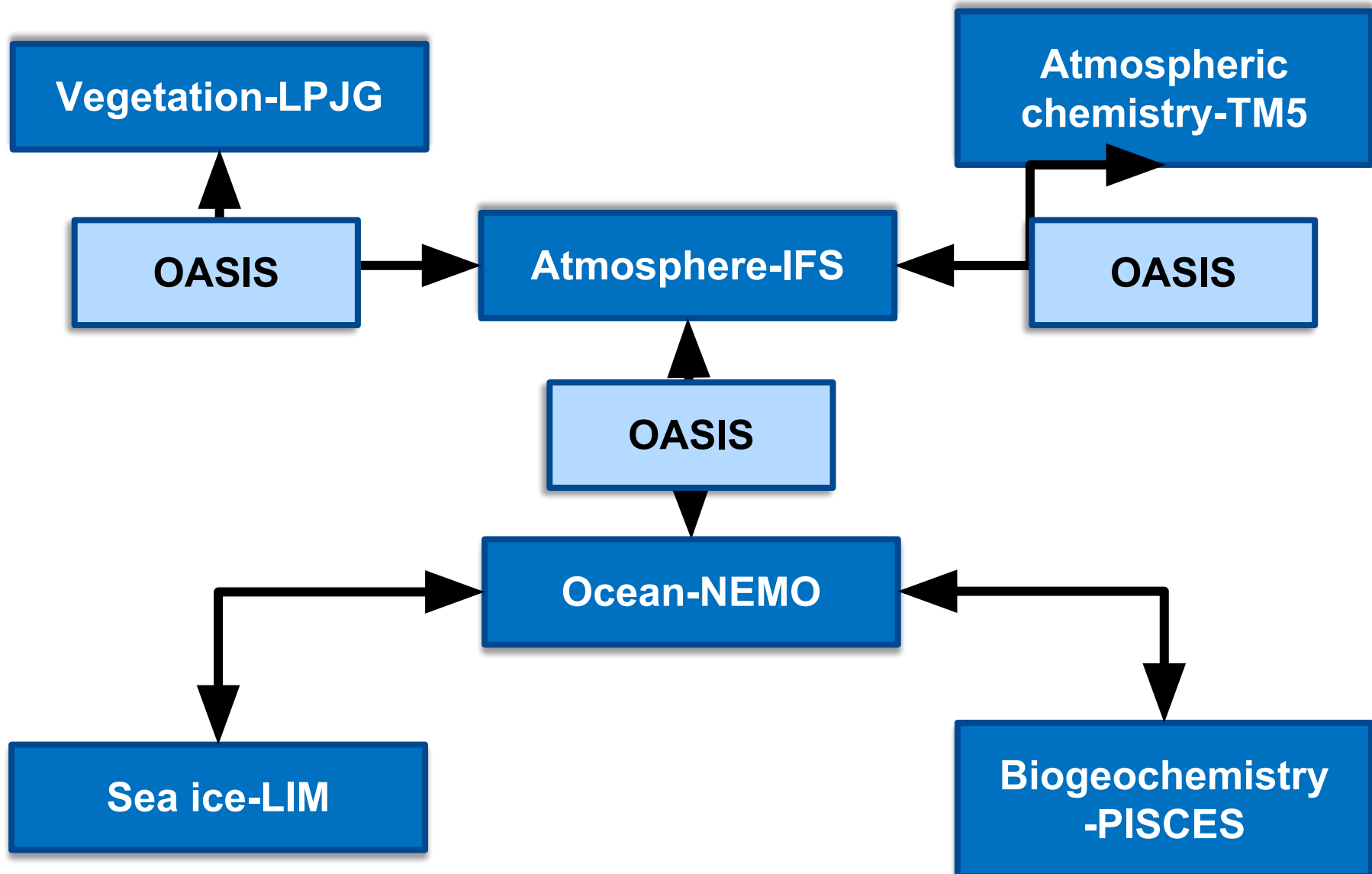


## Data and Diagnostics Team

- Big Data in Earth Sciences
- Provision of internal and external data services
- Visualization

- EC-Earth climate model
  - IFS for atmosphere (ECMWF)
  - NEMO (IPSL) for ocean
- EC-Earth is developed as part of a Europe-wide consortium thus promoting international cooperation
- Scope: develop a fully atmosphere ocean land biosphere model usable for problems encompassing from seasonal to decadal climate prediction to climate change projections and paleoclimate simulations.





# Running a decadal experiment



Barcelona  
Supercomputing  
Center  
Centro Nacional de Supercomputación

EXCELENCIA  
SEVERO  
OCHOA



1981

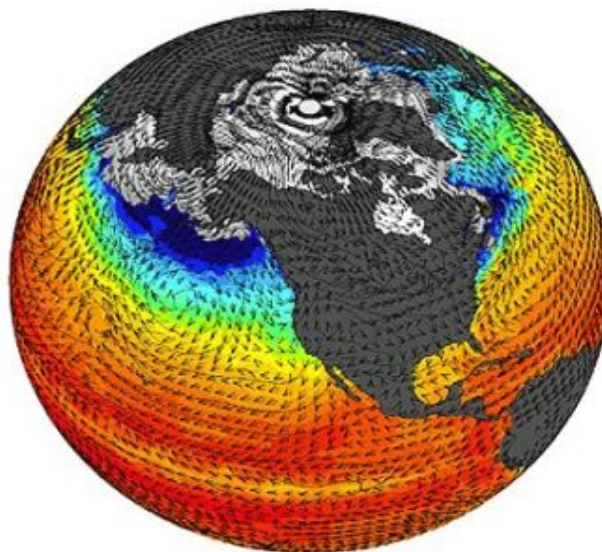
1995

1st Jan

1st Jan

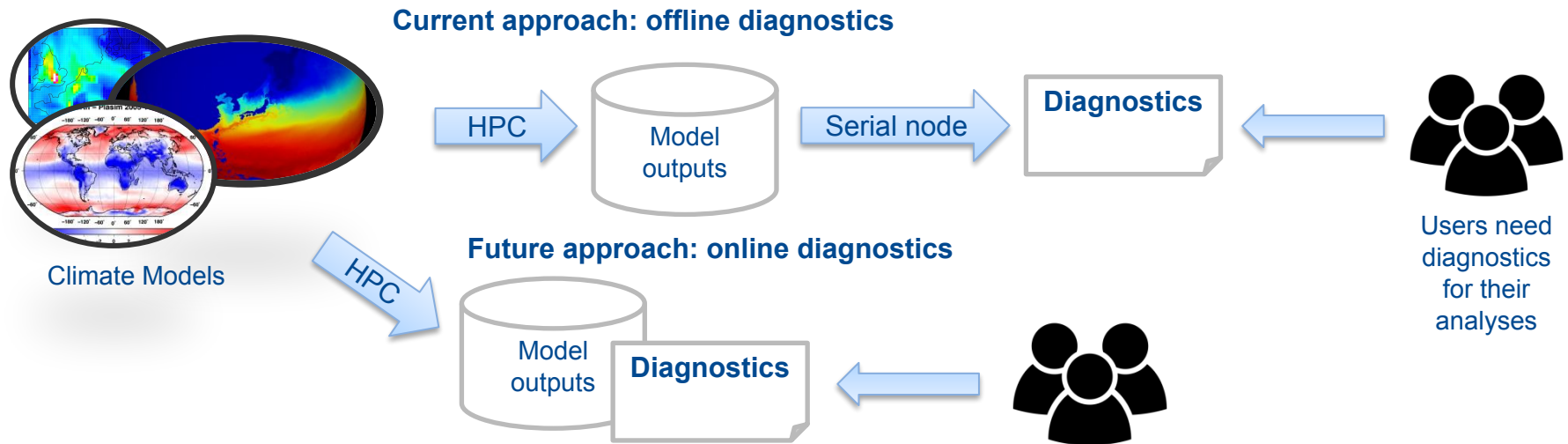
10 years

N members

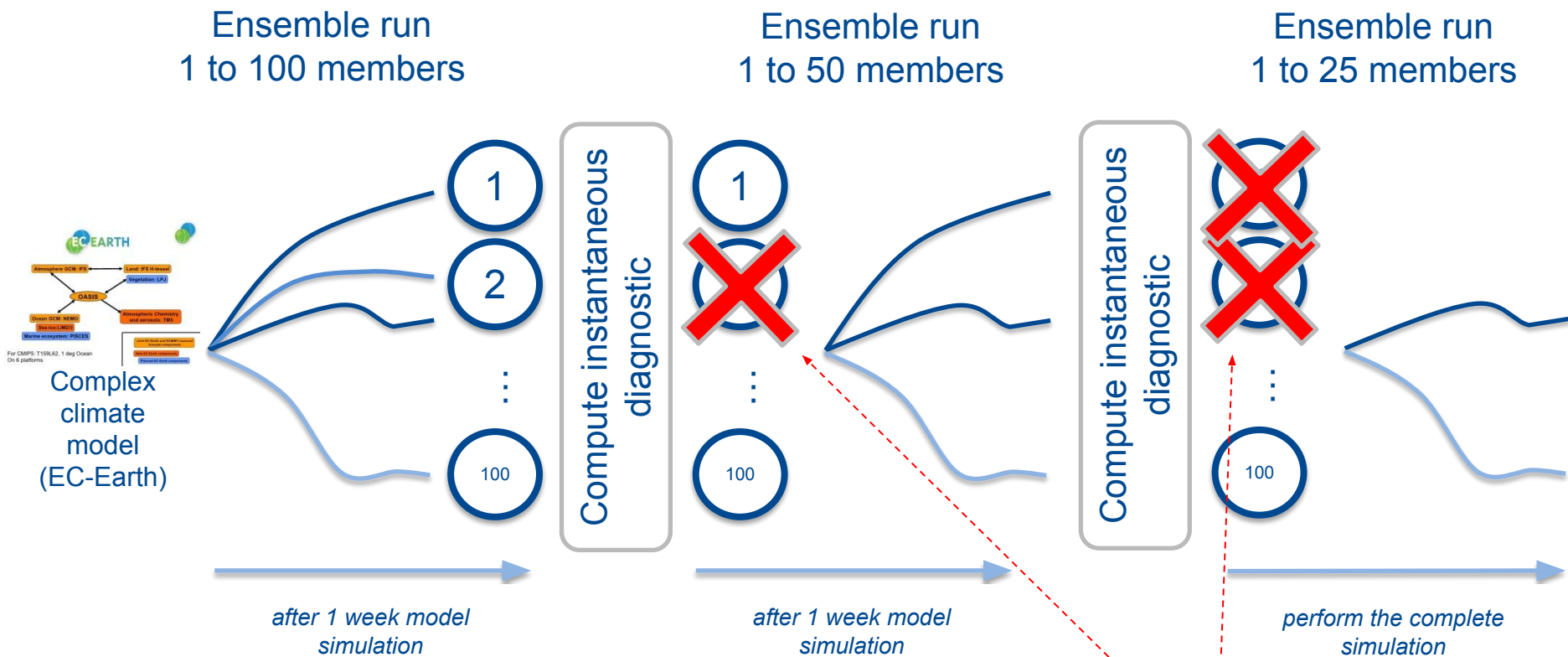




- Workflow manager, monitoring simulation, pre- and post-processing, diagnostics and visualization
  - Operation modes for limited computational resources in decadal climate predictions:
    - N members x 10 years: Currently
    - Other future options: varying the number of members as the forecast time progresses
      - 5N X 1 year
      - $3/2$  N x 1 year
      - N x 1 year
      - $N/2$  x 3 years
      - $N/4$  x 4 years
- Trimming large ensemble simulations as the experiment progresses.  
That requires devoted metrics to optimise ensemble characteristics and speed up selection..



- Diagnostics computed with the Analytics as a Service paradigm
  - Diagnostics online (during model run)
  - Reduced data traffic
  - Diagnostics possible on the computing nodes
  - New diagnostics (data mining of extremes) possible
  - The user gets the results faster → crucial to adapt to climate change and to develop climate services (public and private)



## Key Challenge

How can we perform this computation “instantaneously”?

- Select a right diagnostic (climate science)
- Work in XIOS I/O server and disk technologies (computer science)
- Using PyCompSs to orchestrate the workflow

- Programming Model
  - Sequential programming
  - Portability
  - Standard programming languages
  - APIs
- Runtime system: transparent framework
  - Job & data, resources Management
  - Data synchronization
  - Task scheduling and dependency analysis
- Set of tools
  - Development, execution monitoring and post-performance analysis

For more information, <https://pypi.python.org/pypi/pycompss>

- Ensemble simulation for uncertainty characterization
- Adaptive Ensemble Selection allows to make optimum use of the limited computational resources and tries to prune in an optimal way ensembles from climate models
- PyCompss is a versatile and robust platform, which provides a transparent way to orchestrate the simulation of big ensembles
- Tools, performance and diagnostic selection are key topics to address successfully (and at the same time) the generation and the post-processing of such simulations

# Questions



**Barcelona  
Supercomputing  
Center**  
*Centro Nacional de Supercomputación*



0 1 0 1 0  
1 0 1 0 1  
0 1 0 1 0  
1 0 1 0 1  
0 1 0 1 0  
1 0 1 0 1  
0 1 0 1 0  
1 0 1 0 1  
0 1 0 1 0  
1 0 1 0 1  
0 1 0 1 0  
1 0 1 0 1  
0 1 0 1 0



**Barcelona  
Supercomputing  
Center**

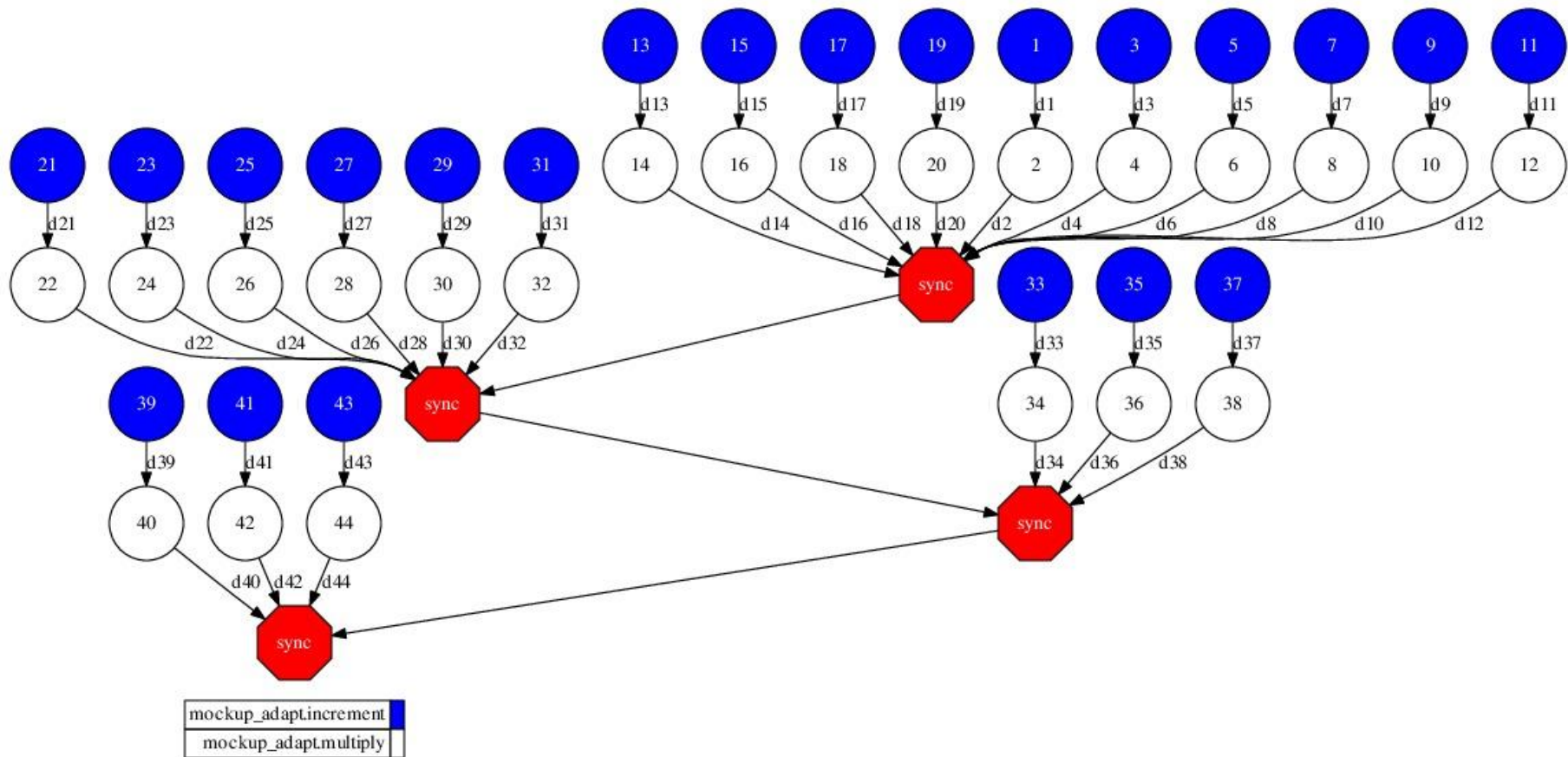
*Centro Nacional de Supercomputación*



# Thank you!

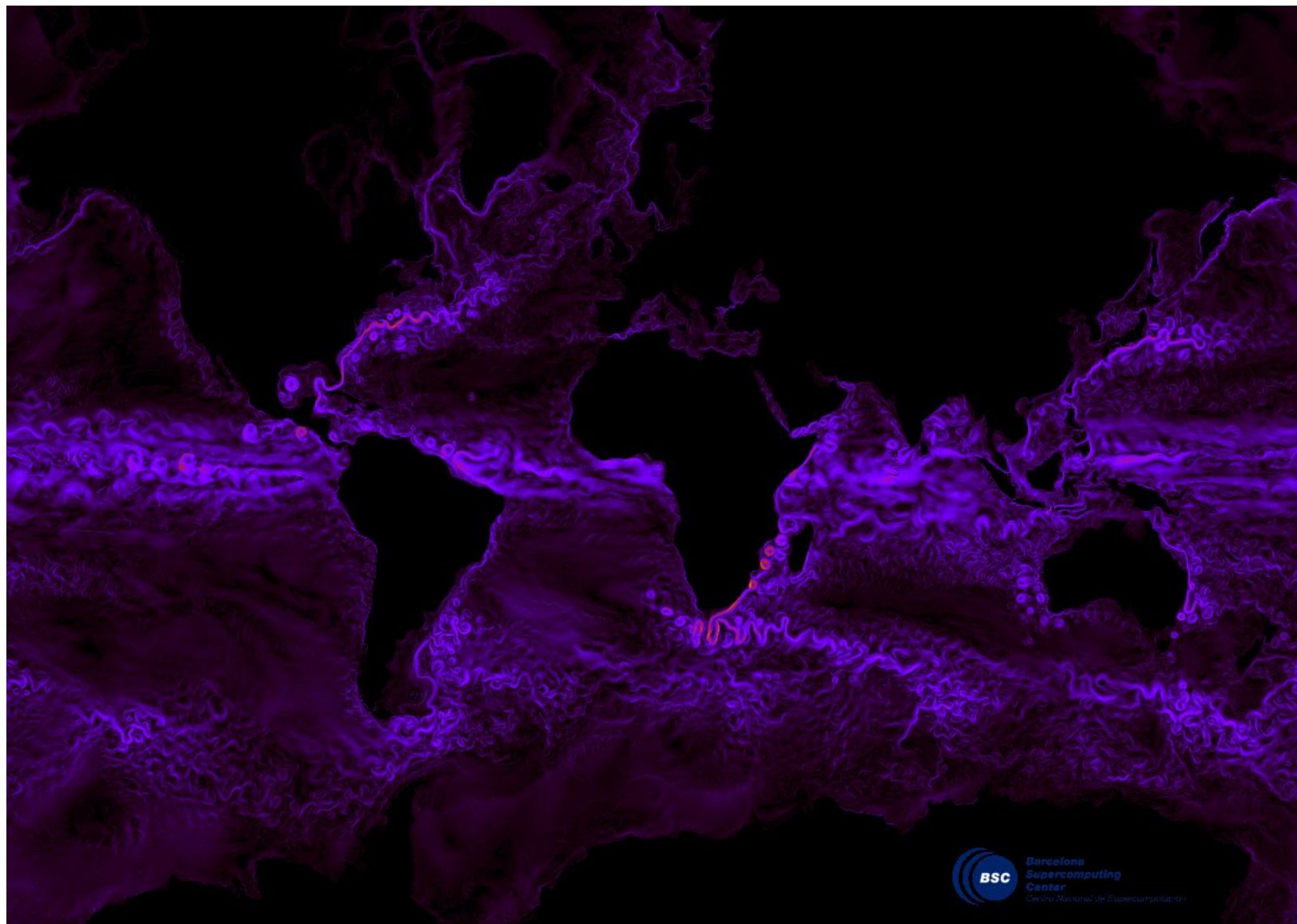
For further information please contact  
[alicia.sanchez@bsc.es](mailto:alicia.sanchez@bsc.es)

# PyCompss MockUp selection





- Diagnostics are computed after simulation
  - Only a subset of variables is saved
- Diagnostics generation is mainly sequential
- Different languages and technologies:
  - Fortran codes
  - Bash scripts using tools like CDO or NCO
  - Python scripts using modules like numpy, iris, pycdo...
- We use fat nodes with a big amount of memory to compute these diagnostics.



*Ocean currents T1279-ORCA12 EC-Earth run (courtesy L. Brodeau)*

- Autosubmit

- A versatile tool to manage Weather and Climate Experiments in diverse Supercomputing Environments
- <https://pypi.python.org/pypi/autosubmit>

